Coarticulation in Second Language Spanish:

Learning to Connect Sounds in a Second Language

Undergraduate Research Thesis

Presented in partial fulfillment of the requirements for graduation with *honors research distinction* in the undergraduate colleges of The Ohio State University

by

Rachel Meyer

Autumn 2020

Project Advisor: Dr. Rebeka Campos-Astorkiza, Department of Spanish and Portuguese

1. Introduction

Coarticulation is a universal feature of language, but it has rarely been studied in a second-language context. While there has been substantive research on acquisition of various sound categories in both English and Spanish, there has been little on the acquisition of language-specific coarticulation, and very little on coarticulation in Spanish at all. Due to the lack of research on coarticulation, we do not know what effect it has on second language acquisition. Different types and degrees of coarticulation can be language-specific features, and therefore must be learned as part of successful second language phonological acquisition.

This study examines how consonant-vowel coarticulation changes between L1 English and L2 Spanish, with a focus on changes over the course of a semester of pronunciation training. Specifically, the influence of voiceless stops on vowels in L2 Spanish will be explored. By looking at the formant values of vowels adjacent to each voiceless stop, I hope to show language-specific patterns of coarticulation. I will also evaluate two models of coarticulation, the feature-spreading model and the degree of articulatory constraint model. Because English and Spanish voiceless stops have different voice onset times, this study also looks at the impact of differing voice onset time durations on coarticulatory patterns. With data from the beginning and end of a semester of pronunciation training, I will also look for development in coarticulation as more knowledge of Spanish phonology is acquired by the participants.

2. Background

Coarticulation occurs when the gestures between speech sounds overlap. This overlap occurs because the articulators are moving continuously, generating intermediate productions as they shift between sounds. Coarticulation can anticipate features from a following sound or carry over features from a preceding sound. Because coarticulation varies among languages (Farnetani and Recasens 1997), L2 learners might need to acquire new coarticulatory patterns.

One of the first and most comprehensive general studies on coarticulation in English was conducted by Stevens and House (1963). They hypothesized that coarticulation occurs when the articulators cannot move quickly enough between sounds and do not reach the target position. They found that there were strong effects on vowels from surrounding consonants, but this effect varied by vowel. They also found that place, manner, and voicing of the consonant affected the vowel, although F2 experienced greater shifts than F1. The identity of the vowel also mattered for coarticulation. Front vowels shifted more in response to both labial and postdental consonants, while back vowels shifted more in response to postdental consonants only.

Fowler and Saltzman (1993) hypothesized that the mechanism of coarticulation was based on “blending strength.” They conceptualized each class of sounds as having a unique susceptibility to coarticulation. This susceptibility is inversely correlated with sonority, such that stops have the most blending strength and vowels have the least. Based on this metric, stops would be more likely to impose their features on sounds with less blending strength, and to resist changing features in response to adjacent sounds. Conversely, vowels are likely to accept features from other sounds, and not very likely to impose features on other sounds.

There has not been much research on coarticulation in L1 Spanish specifically. However, one study by Recasens (1987) compared Spanish and Catalan coarticulation in VCV sequences. Participants read a series of sentences with the target sounds. The target sounds were laterals, rhotics, and voiced stops, and vowels in V-to-V and V-to-C sequences in both languages. He found that carryover coarticulation is more dependent on articulatory constraints than anticipatory coarticulation. Anticipatory coarticulation was more of a temporal effect than an articulatory one.

Solé and Ohala (1991) looked at nasal coarticulation in English and Spanish. More precisely, they explore the hypothesis that in English, in words like "hand," nasalization is a phonological process, and the vowel is always nasalized, while in Spanish, nasalization can occur when moving between sounds, but it is not phonological. Solé and Ohala recorded speakers of both languages saying words with -VN syllables at several speech rates. They found that the percentage of the vowel duration that was nasalized varied in Spanish, but the entire vowel was nasalized in English regardless of speech rate. They interpreted this result as evidence for merely phonetic nasalization in Spanish. Spanish vowels are nasalized because of "physiological time constraints." The variable length of nasalization in English is evidence that it is a phonological process, the expected form of the vowel in a -VN environment. This study demonstrated that coarticulation can vary across languages, even when similar sounds are involved. It also demonstrated that coarticulation can be phonological, as in English nasal vowels, or phonetic, as in Spanish nasal vowels.

There has been very little study on coarticulation in a second language Spanish context. However, one study, by Solon (2018) explored vowel-/l/ coarticulation at various levels of Spanish proficiency. In English, a syllable-final /l/ is usually velarized (dark), while in Spanish it is always non-velar (light). It has been shown cross-linguistically that light /l/ is less coarticulatorily resistant than dark /ɫ/ (Recasens and Farnetani 1990). The level of coarticulatory resistance varies among languages, though (Recasens 2012a). Solon predicted that syllable-final /l/ would be subject to less coarticulation in more experienced students' productions. Students with less Spanish experience would be more likely to use a darker, more coarticulatory resistant

/l/ before they learned the lighter, less resistant /l/ of Spanish. Solon looked at undergraduate students with 1-4 years of Spanish study, graduate students, and native speakers. She elicited data through spontaneous and sentence- and word list-reading tasks. She found that L2 /l/ productions were in between L1 English and L1 Spanish in terms of lightness. She also found that L2 Spanish productions were intermediate in terms of coarticulatory resistance. This was visible through the difference in F2 of the /l/ when it was preceded by a front vowel versus a back vowel. Her finding varied as a function of experience, with more and more differentiation between an /l/ preceded by a front vowel and an /l/ preceded by a back vowel as students learned more Spanish. She concluded that coarticulatory effects on /l/ develop with experience.

There are two primary models of coarticulation that will be examined in this study. The first, the feature spreading model, predicts that a segment specified for a particular feature will "spread" that feature to preceding segments (Hammarberg 1976). In this model, the "feature" would be place of articulation (POA), which would spread from stops to vowels. Since vowels are not specified for frontness or backness as strictly as stops are for place of articulation, vowels can shift in position according to the stop without any loss in identifiability (Keating 1988). The place of articulation feature will spread backwards over preceding sounds until it reaches a sound specified for place of articulation, at which point it will be blocked by that feature. This process is called "look-ahead" because the articulators are constantly looking forward to specified features that will be relevant later. Because of look-ahead, only anticipatory coarticulation is predicted in the feature-spreading model, not carryover coarticulation (Volenec 2015).

The second model of coarticulation is the degree of articulatory constraint model (DAC) (Recasens et al., 1997). This model posits that each phoneme has a unique susceptibility to coarticulation, based on its coarticulatory resistance, which is in turn based on its features.

Higher DAC values designate more constrained or resistant sounds, which exert more carryover coarticulation and less anticipatory coarticulation. Higher DAC values also win out when there are competing features. In this model, stops have a higher DAC value than vowels. For example, in a syllable like /po/, the /p/ has a very fronted place of articulation, at the lips, while the /o/ is a back vowel. Since /p/ has a stronger DAC than /o/, coarticulation would occur by the /o/ becoming more fronted, rather than the /p/ becoming more backed (Recasens 1987). This model has separate predictions for carryover and anticipatory coarticulation. These predictions are mirrors of each other. For example, if /p/ exhibits strong carryover coarticulation on vowels, it will exhibit weaker anticipatory coarticulation.

In both models, the effects of stop-vowel coarticulation would be reflected in the vowel's formant frequencies. The formants are acoustic characteristics that correlate with the articulatory features of height and backness. The formants help identify where the vowel is being produced in the mouth, and can demonstrate coarticulation by their values. The first formant (F1) corresponds to vowel height, which changes based on height of the jaw during vowel production. Vowels produced with high jaws, like /i/, have a low F1. Vowels produced with low, open jaws, like /a/, have a high F1. In the feature spreading model, we expect the jaw heights of following stop consonants to influence vowel height. According to Recasens (2012b), /t/ has the highest jaw or tightest closure during production. In the feature spreading model, a higher jaw means that F1 will be lowest for vowels adjacent to /t/, with the next highest being vowels adjacent to /k/, and finally /p/.

The second formant (F2) corresponds to vowel backness. Front vowels have high F2, like /e/, and back vowels have low F2, like /o/. In the feature spreading model, the place of articulation of the stop consonant will affect the backness of the vowel. Since /k/ is the most backed stop, the vowel will be pulled farthest back when adjacent to /k/, having a lower F2. The F2 value will increase as the adjacent consonant is farther front, so vowels adjacent to /t/ will have the next highest F2, then /p/.

In the DAC model, /t/ is the most constrained voiceless stop because it requires the most tongue precision. Since it does not involve the tongue, /p/ is the least constrained, and /k/ falls in between. The DAC predicts that more constrained phonemes show more carryover coarticulation, and less anticipatory coarticulation. Therefore, there are different predictions for carryover and anticipatory coarticulation in this model.

In carryover coarticulation, vowels following a more constrained sound will have a lower F1 due to carryover of the higher jaw in the preceding stop. They will have a higher F2 due to carryover from the preceding stop place of articulation. For anticipatory coarticulation, the lowest F1 values will result from vowels preceding the least constrained stops. Vowels preceding less constrained sounds will also have the higher F2 values. These predictions are summarized in Table 1 below. I will evaluate these predictions against my data and determine which, if either, of the models does a better job explaining coarticulation.

	F1		F2	
	Carryover (preceding stop)	Anticipatory (following stop)	Carryover (preceding stop)	Anticipatory (following stop)
Feature Spreading	p=t=k	t<k<p	p=t=k	k<t<p
Degree of Articulatory Constraint	t<k<p	p<k<t	p<k<t	t<k<p

Table 1. Summary of predicted value of formants in vowels adjacent to the noted consonant according to the two models of coarticulation considered in this study.

As this study discusses voiceless stops, it is important to note that Spanish and English differ in their stop inventories. Spanish and English have the same three voiceless stop phonemes: /p/, /t/, and /k/, but voiceless stops in English in most contexts are produced with a markedly longer voice onset time (VOT) than Spanish stops (Hualde 2005). English voiceless stops are described as aspirated because of their long VOT, while Spanish voiceless stops are unaspirated. The failure to accommodate to the new, shorter VOT is one contributor to foreign accent in Spanish for native English-speaking learners. Additionally, Spanish has only five vowels where English has around eleven, although all five Spanish vowels have close matches in English, and English speakers with no Spanish experience are capable of discriminating Spanish vowels (Menke and Face 2010).

Because VOT differentiates English and Spanish voiceless stops, it is possible that the VOT might play a role in coarticulation. With a longer VOT, there is more distance between the closure for the stop, at a specific place of articulation, and the vowel. This might reduce coarticulation because there is more separation between articulatory gestures. In English, there is more time between the stop burst and the vowel, so there might be less of an effect of coarticulation than in Spanish.

This study will examine how coarticulation in second language Spanish changes over time by looking at data from students at the beginning and end of a Spanish pronunciation class. In the context of the two theories of coarticulation discussed above, I will examine how coarticulation varies between English and L2 Spanish. I will also examine how aspiration interacts with coarticulation, since it is the main difference between the voiceless stops in the two languages.

Research Questions:

- How does coarticulation vary between English and L2 Spanish?
- How does coarticulation change over the course of a semester?

- How does aspiration interact with coarticulation?

3. Methodology

3.1 Participants and Stimuli

The data analyzed for this study come from See Your Speech, an interactive pronunciation project. See Your Speech originated as a project examining dialectal change in college students as they came to Ohio State from the various dialect regions in Ohio and across the country (Wanjema et al., 2013). This project expanded to the language departments at Ohio State, and now is used to combine research and teaching on L2 pronunciation and phonological acquisition.

All of the data comes from a module presented to students in SPAN 3404, a college-level Spanish pronunciation class required for both majors and minors in Spanish. Using a web-based interface, students record themselves reading word lists in both English and Spanish at the beginning and end of the semester. They receive instantaneous feedback on their pronunciation in both languages, including vowel plots. This study analyzes both the English and Spanish data from 53 students. All students were native American English speakers and began learning Spanish between the ages of 12-14.

The stimuli are presented to the students one by one, in ten word groups, as they complete the module. There are 71 words in Spanish and 60 words in English. The stimuli target a variety of sounds in both languages, including stops and vowels. However, they were designed to look at vowel, stop, and lateral acquisition in Spanish, and regional variation in English. This presents some limitations, since the stimuli were not specifically designed to study coarticulation. In the Spanish wordlist, voiceless stops appear in both intervocalic and initial positions. The vowels may be stressed on unstressed. The comparison word list in English includes stops primarily in initial position, in monosyllables.

The English stimuli used for this project consist of 16 words, and the Spanish data is made up of 17 words. All the words have the same structure in each language. The English stimuli have the structure consonant vowel consonant (CVC), with the first consonant being a voiceless stop. The Spanish words have a CVCV structure, and both Cs are voiceless stops. The stimuli were specifically chosen to be representative of English and Spanish vowel and stop inventories (Table 1).

English		**Spanish**	
caught	peace	capa ("cape")	pito ("whistle")
coat	pet	capó ("she/he neutered")	pitó ("she/he whistled")
code	pout	coca ("coke")	taco ("taco")
cot	put	papa ("potato")	tapo ("I cover")
cud	taught	pato ("duck")	tapó ("she/he covered")
cuss	test	peco ("I sin")	toco ("I touch")
paid	tooth	pecó ("she/he sinned")	tocó ("she/he touched")
pass	toyed	pico ("peak")	tope ("limit")
		picó ("it itched")	

Table 1. Stimuli lists in English and Spanish. Words were chosen to represent the voiceless stop and vowel inventories of each language.

3.2 Acoustic Analysis

The sounds of interest were segmented by hand in Praat (Boersma and Weenink 2005), a computer program to carry out acoustic analysis of recordings. The target vowels' F1 and F2 were measured, and the target consonants' voice onset time duration was measured. Praat visualizes the sounds in both waveform and spectrogram forms (Figure 1). The dark bands on the spectrogram represent the formants, with the darkness corresponding with the amplitude of the sound (marked in yellow on the spectrogram). Vowels in general have a higher amplitude than consonants. Changes in amplitude are used to establish onset and offset of vowels. This is paired with the waveform through the onset and offset of regular vibration cycles. Stop releases typically begin with one or more bursts, observable in the waveform. After the burst there is a period of aspiration, which makes up the voice onset time (VOT). The VOT is the time between the burst and the onset of voicing for the following vowel. The end of the VOT corresponds to an increase in amplitude and the onset of voicing.

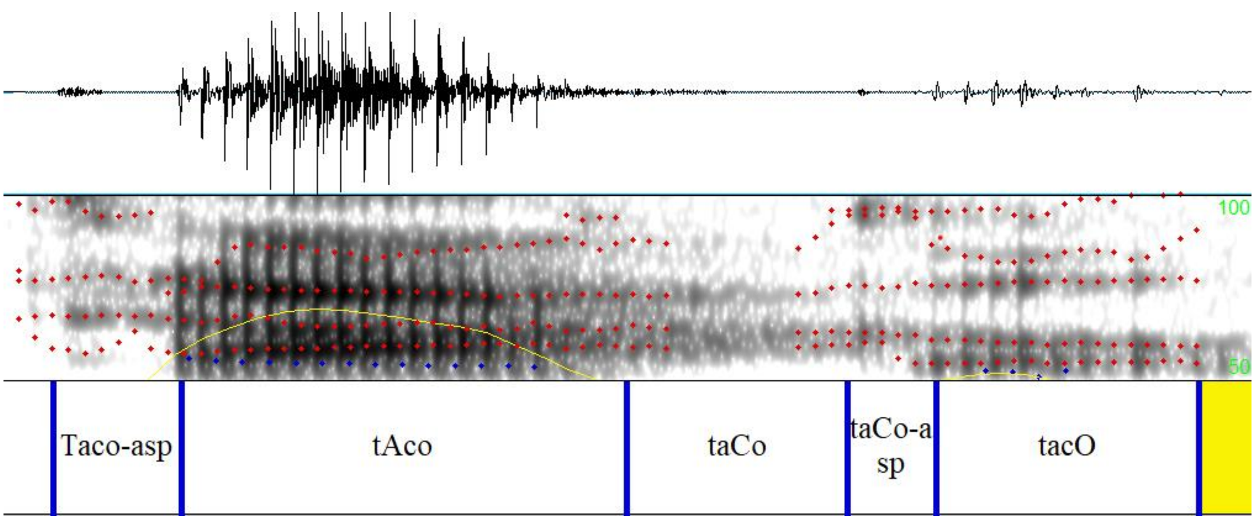

Figure 1. The segmentation of the word "taco." The red dots mark the formant frequencies, and the yellow line is the intensity of the sound. In the bottom tier, the capital

letter refers to the sound being produced at that time. The aspiration (VOT) of the stop is also marked.

After the stops and vowels were segmented, the VOT duration (for stops) and formants (for vowels) were measured. The first two formant frequencies were measured at three timepoints across the vowels' duration: at 20%, 50%, and 80% of the length of the vowel. The length of the voice onset time, from the stop burst to the vowel onset was measured for consonants. Each token was measured at the beginning and at the end of the semester. The measurements will be examined for change across the duration of the vowel, resulting from the influence of surrounding sounds, and over the course of the semester. The duration of the voice onset time of the stop was also considered for potential effects.

3.3 Data Analysis

Once all the segmentation and measurements were complete, the total number of tokens was 3564 (1320 English, 2344 Spanish). Tokens were excluded for poor sound quality or mispronunciation, but tokens with the stress on the wrong syllable were still included. Several factors' influence on vowel production were explored: place of articulation of the consonant, language (English or Spanish), and timepoint (beginning or end of the semester). Interactions between these factors were also examined. The influence of the preceding or the following stop on vowels was considered separately, although this division only applies to the Spanish data. Due to the nature of the English wordlist, the stimuli only had preceding stops. Descriptive statistics were used for VOT and the formants (F1 and F2), by stop and vowel combinations. Linear regression was used to determine the significance of the factors and interactions mentioned above on the F1 and F2 values. Participant was included as a random factor. Models were built for each data subset at each of the vowel durations, i.e. 20%, 50% and 80% (see more details in the results section) and compared using ANOVA to determine the best-fit model. In addition, pairwise comparisons helped explore significant interactions. The statistical significance level was set at p-value < 0.05. See Appendix A for tables with the best fit models from the linear regression analysis.

4. Results

Data limitations affected the way the results had to be examined. Because each vowel did not occur in the stimuli equally with each consonant, it is necessary to look at separate vowel groupings individually. Pooling all the data results in a confound between vowel quality and stop consonant, making large effects appear when in reality it is just a disproportionate number of stimuli with a particular CV combination. Additionally, the different formants of each vowel make it difficult to look at all vowels together. To address these issues, I sorted the vowels into categories via height (for F1) and backness (for F2). The categories with enough data from all three places of articulation that will be examined are back vowels and central vowels (for F2) and mid vowels and low vowels (for F1). For high vowels and front vowels, there was not enough data at all three places of articulation to merit looking at them individually.

Furthermore, because of the structure of the stimuli, there is only data on the preceding consonant for the English data, as mentioned earlier. In this section, I will present the results for coarticulation in English and Spanish vowels preceded by stop consonants and Spanish vowels followed by stop consonants. The statistical analyses are carried out for these data subsets separately. Finally, I will look at the impact of VOT on coarticulation.

For vowels preceded by a stop consonant, we are expecting the effects of the consonant to be strongest at the beginning of the word, which is closest to the consonant. The prediction is that there will be differences at 20% of the vowel duration, but these will decrease or be non-existent by 50% or 80% of the vowel duration. Conversely, for vowels followed by a stop consonant, we expect to see a bigger effect at 80% of the vowel duration, which decreases backwards over the duration of the vowel, with very little effect at 50% or 20% of vowel duration.

4.1 Results for F2

4.1.1 Back Vowels

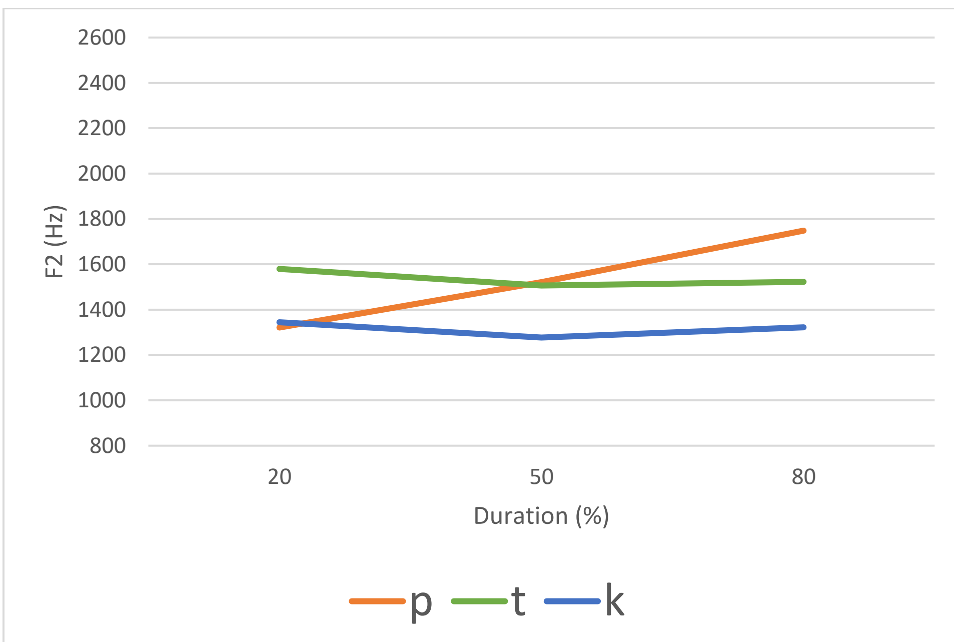

Figure 2. Average F2 of English back vowels preceded by voiceless stops.

In English, vowels preceded by /k/ have a lower F2 than vowels preceded by /t/ throughout their duration (Figure 2). Both of these groups have the highest F2 at 20% of vowel duration, then fall towards the middle of the vowel and stay fairly constant to the end of the vowel. Vowels preceded by /k/ have a slightly higher F2 at 80% of vowel duration than at 50%. Vowels preceded by /p/ have a different trend. The F2 is lowest at the beginning of the vowel and rises over the course of the vowel. F2 differences based on the preceding consonant can be seen across the entire vowel.

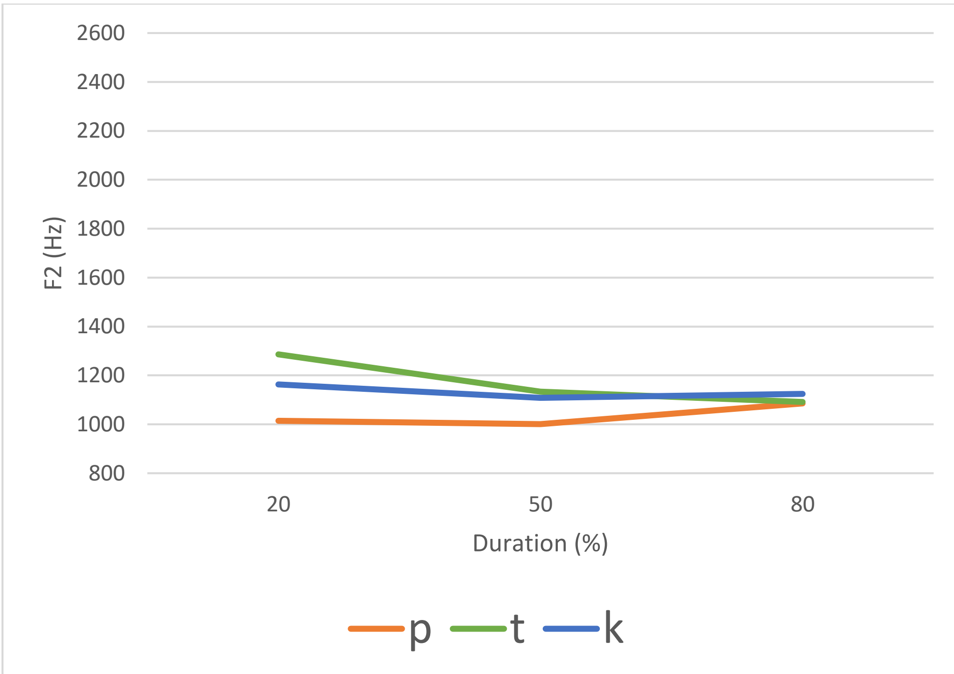

Figure 3. Average F2 of Spanish back vowels preceded by voiceless stops.

In Spanish, the F2 is lower on average than in English. This is expected since the back vowels are produced more backed than in English. In the Midland dialect of American English that many of the participants speak, /u/ and /o/ are centralized, resulting in a higher F2.

At 20%, we see large differences between Spanish vowels preceded by each voiceless stop (Figure 3). Vowels preceded by /p/ have the lowest F2, then vowels preceded by /k/, and finally vowels preceded by /t/ have the highest F2. The F2 for vowels preceded by /k/ stays roughly the same over the course of the vowel, but vowels preceded by /t/ have a lower F2, equivalent to that of vowels preceded by /k/ at 50% of the vowel. These two groups of vowels have approximately the same F2 at the end of the vowel as well. Vowels preceded by /p/ have about the same F2 at 50% vowel duration as they did at 20%, but the F2 rises slightly to meet /t/ and /k/ by the end of the vowel. In summary, we see broad differences based on place of articulation at the beginning of the vowel, but similar F2 values at the end of the vowels.

When looking at the statistical results for the preceding consonant dataset, place of articulation was significant for the F2 at the three measured durations. However, at 50% the difference in F2 for /t/ and /k/ was not significant, and at 80% only the difference between /k/ and /p/ was significant. Language was also significant at all three duration measurements, which is expected because of formant differences between languages in terms of their vowels, as I mentioned earlier. Finally, there was a significant interaction between place of articulation and language at the three durations.

Examining the pairwise comparisons, all comparisons are significant at 20%, except for English /p/ and /k/. This is visible in Figure 2, where vowels preceded by /p/ and /k/ have the same F2 at 20% duration. At 50% duration, in English vowels preceded by /t/ do not have a different F2 from vowels preceded by /p/. In Spanish at 50% duration, vowels preceded by /k/ do not present a different F2 from vowels preceded by /t/. At 80% duration, none of the Spanish vowels are significantly different from each other.

There is no effect of timepoint at any of the three duration measurements. We do not see an interaction between place of articulation and language, and no three-way interaction between timepoint, place of articulation, and language.

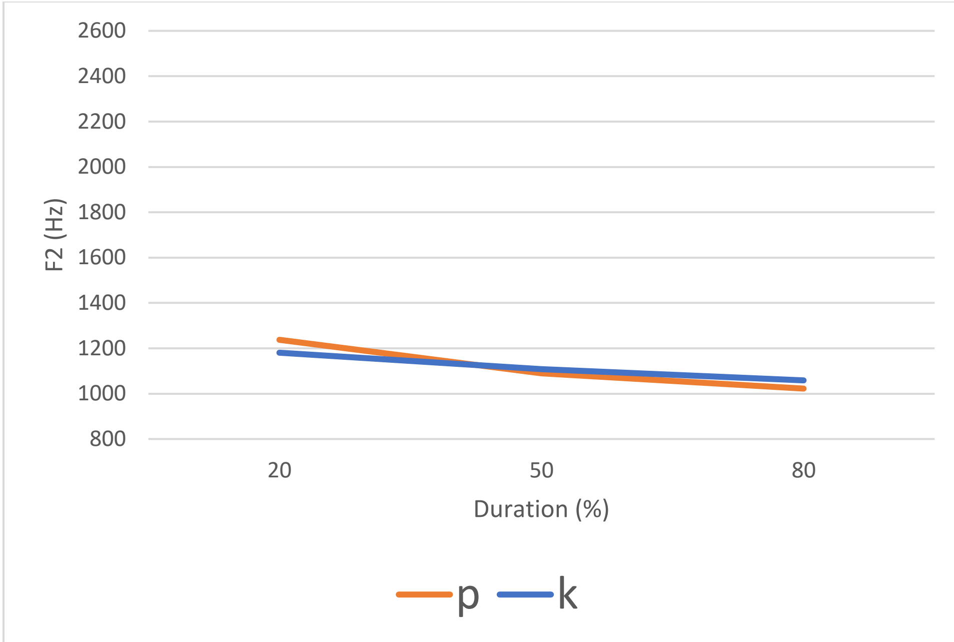

Figure 4. Average F2 of Spanish back vowels followed by voiceless stops.

In the stimuli, back vowels are only ever followed by the stops /p/ and /k/ (Figure 4). Both sets of vowels have approximately the same F2 throughout the vowel. It is highest at the beginning of the vowel and declines towards the middle and end of the vowel.

Place of articulation is not significant here, and timepoint is only significant at 80% of vowel duration. The F2 is higher at the beginning of the semester than at the end of the semester. There is no interaction between place of articulation and timepoint at any point over the duration of the vowel.

4.1.2 Central Vowels

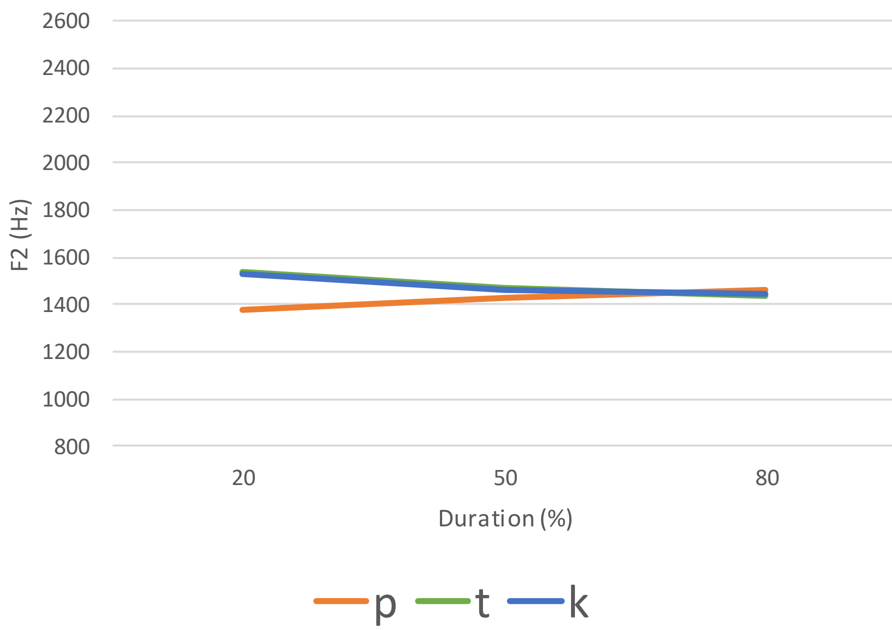

Figure 5. Average F2 of Spanish central vowels preceded by voiceless stops.

In Spanish, central vowels preceded by /p/ begin with a lower F2 than vowels preceded by /t/ or /k/ (Figure 5). The entire set of vowels converge in F2 by 50% duration, and are approximately the same by the end of the vowel. Place of articulation is significant at 20% and 50% duration, although vowels preceded by /t/ are the same as vowels preceded by /k/ at all duration points. In English, the central vowel is only ever preceded by /k/, so no effect of place of articulation on coarticulation can be explored. However, there is an effect of language. There is no effect of timepoint and no interactions between factors.

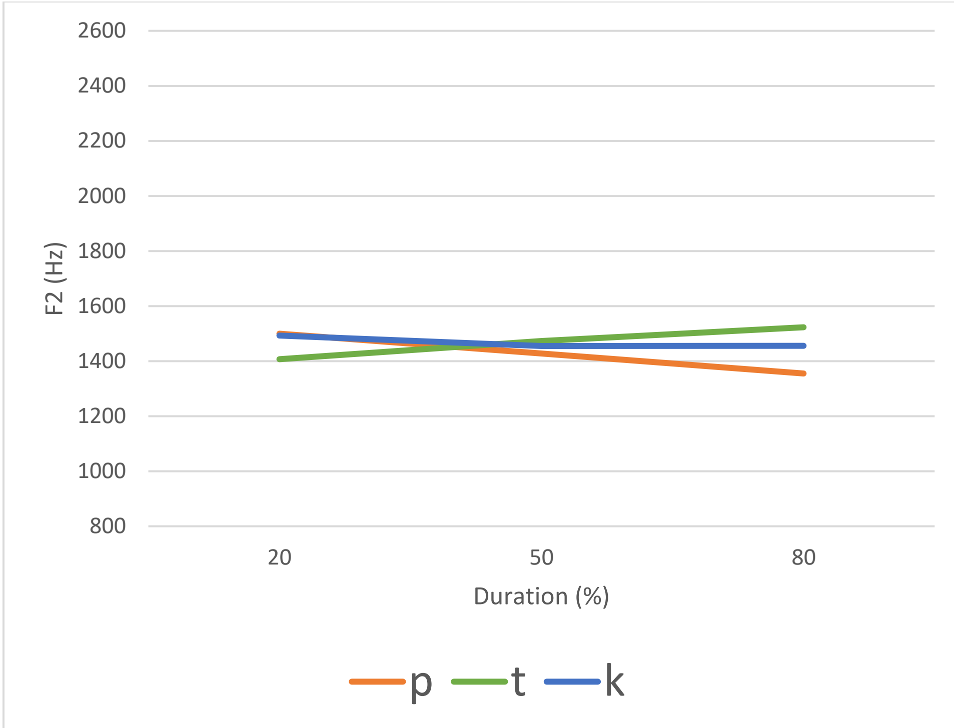

Figure 6. Average F2 of Spanish central vowels followed by voiceless stops.

For central vowels followed by a voiceless stop, there is a bit of a different trajectory. Vowels followed by /p/ or /k/ start out at approximately the same value, and decline over the course of the vowel, although, vowels followed by /p/ drops a bit more steeply (Figure 6). Conversely, vowels followed by /t/ start out lower than vowels followed by /p/ or /k/ and rises over the course of the vowel, having the highest average F2 by the end of the vowel. The biggest F2 differences based on the following consonant are seen at 80% of the vowel.

The statistical results show that place of articulation is significant at 20% and 80%, although vowels followed by /p/ do not have a different F2 from vowels followed by /k/ at either duration. Place of articulation is not significant at 50% vowel duration and timepoint is not significant at any of the durations.

4.2 Results for F1

4.2.1 Mid Vowels

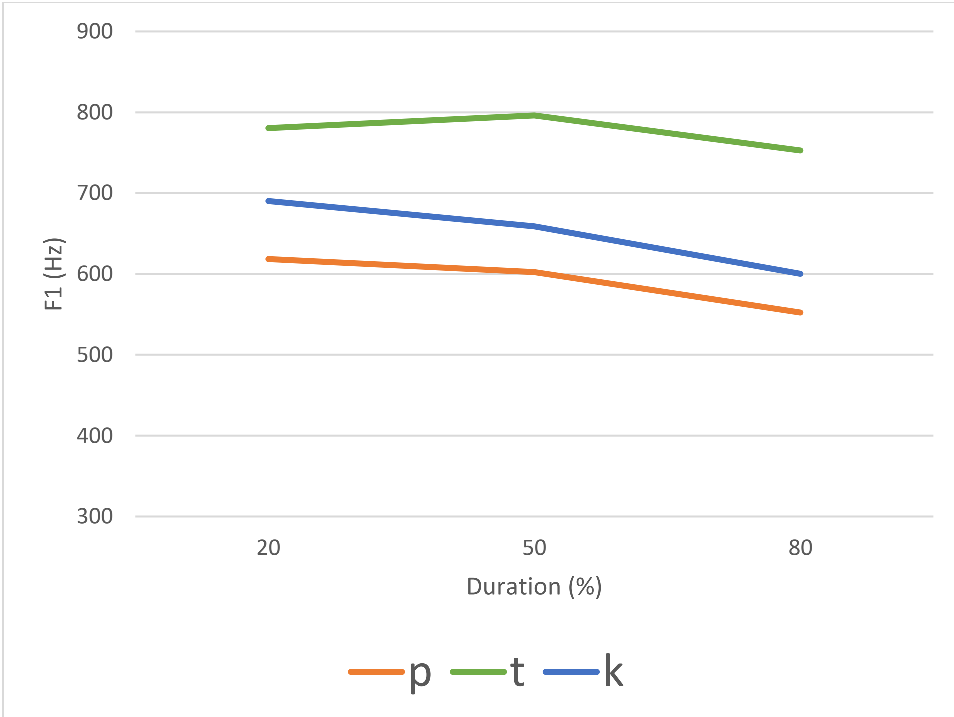

Figure 7. Average F1 of English mid vowels preceded by voiceless stops.

In English, there are large differences in the F1 of mid vowels depending on the preceding voiceless stops (Figure 7). F1 is highest for vowels preceded by /t/, then vowels preceded by /k/, then vowels preceded by /p/. Vowels preceded by /p/ and /k/ decline in F1 over the course of the vowel. The F1 of vowels preceded by /t/ increases slightly towards the middle of the vowel, then drops towards the end of the vowel. The F1 is also markedly higher in vowels preceded by /t/, and there is a fairly large difference between vowels preceded by /k/ and vowels preceded by /p/ as well.

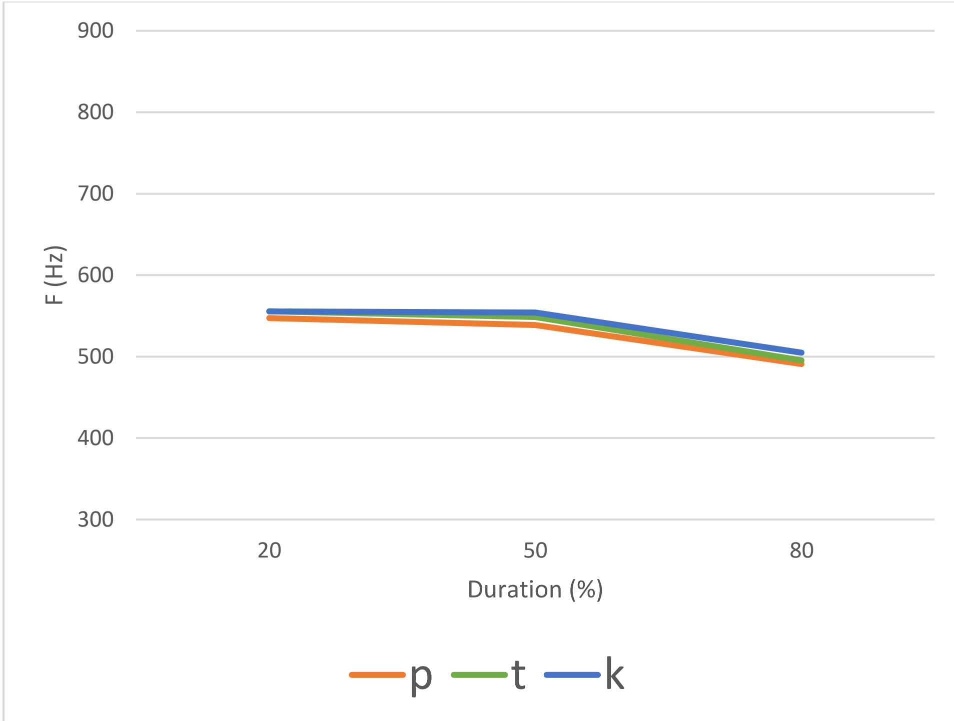

Figure 8. Average F1 of Spanish mid vowels preceded by voiceless stops.

In Spanish, the results look very different. The F1 is lower on average, and all of the vowels' F1s are approximately the same regardless of preceding consonant (Figure 8). The F1 stays the same until about the middle of the vowel, and then drops off towards the end.

When considering the statistical results, there is an overall effect of place of articulation and language. There is also an interaction between place of articulation and language, which shows that there are no differences between places of articulation in Spanish, although there are in English. There is no effect of timepoint.

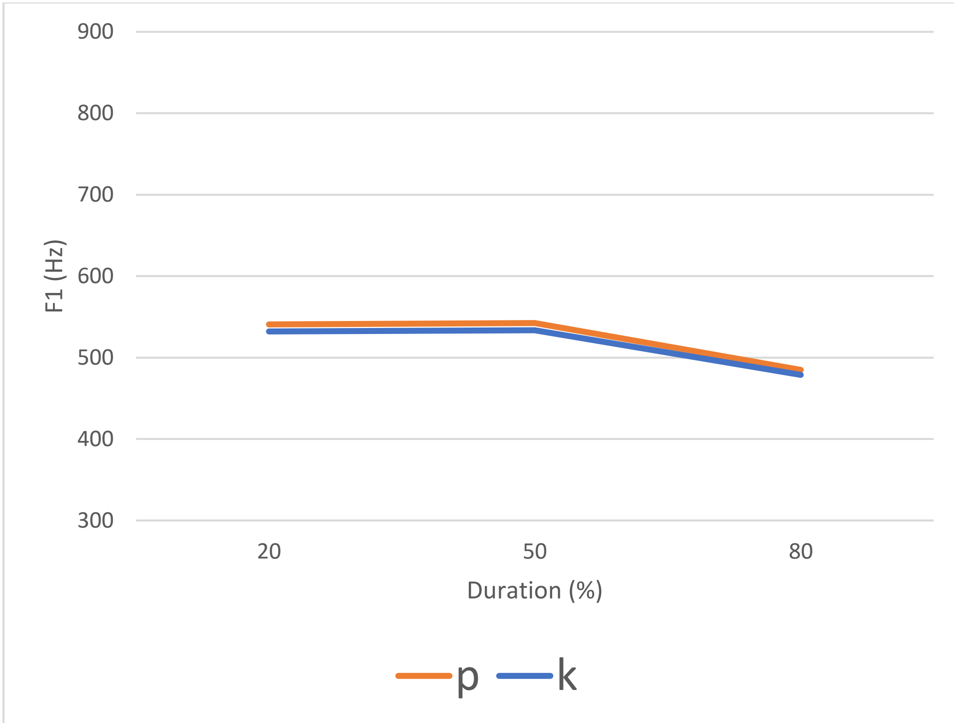

Figure 9. Average F1 of Spanish mid vowels followed by voiceless stops.

The lack of effect of place of articulation in Spanish vowels followed by a voiceless stop is similar to vowels preceded by one (Figure 9). There is a similar trajectory-the F1 declines in the second half of the vowel’s duration. There is an effect of timepoint at 50% and 80%. At the beginning of the semester, the F1 is higher at 50% vowel duration and lower at 80% vowel duration, in comparison to the end of the semester. There is no effect of place of articulation and no interaction between place of articulation and timepoint.

4.2.2 Low Vowels

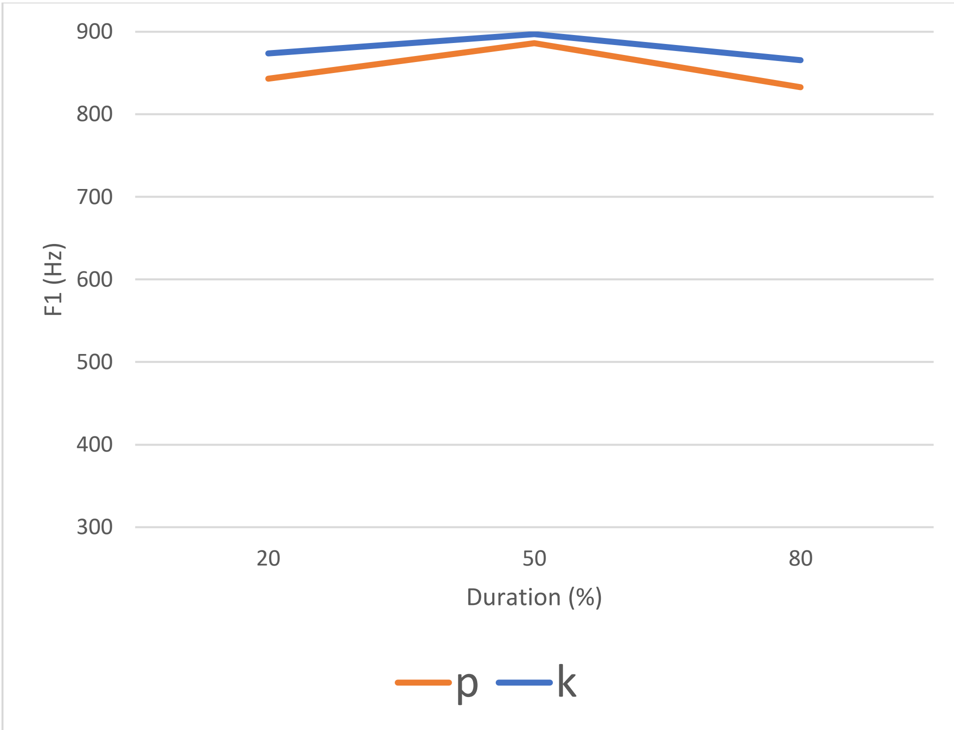

Figure 10. Average F1 of English low vowels preceded by voiceless stops.

In English, low vowels preceded by /p/ have a slightly lower F1 than low vowels preceded by /k/ (Figure 10). The F1 peaks in the middle of the vowel, and is a little bit lower at the beginning and end of the vowel.

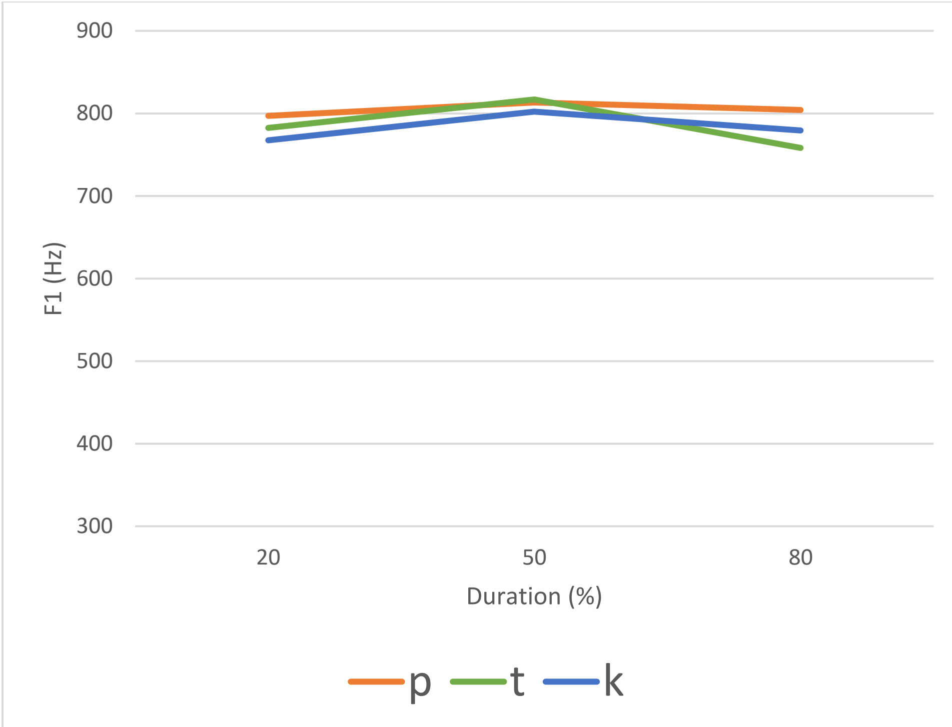

Figure 11. Average F1 of Spanish low vowels preceded by voiceless stops.

In Spanish, the average F1 is lower than in English, but the vowels follow a similar trajectory to English, with the peak of F1 in the middle (Figure 11). Vowels preceded by all three voiceless stops are fairly close together. The statistical results show that at the end of the vowel there is a significant difference between vowels preceded by /p/ and vowels preceded by /t/ Place of articulation is not significant at 20% or 50%. There is also an effect of language, F1 is lower in Spanish than in English. Finally, there is no effect of timepoint.

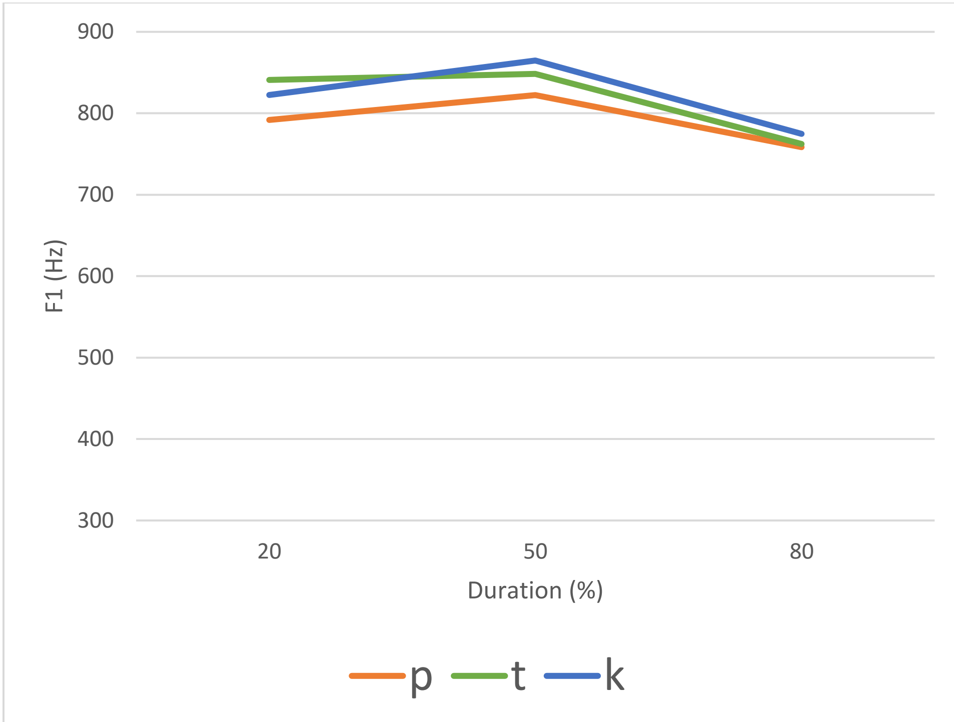

Figure 12. Average F1 of Spanish low vowels followed by a voiceless stop.

For low vowels followed by a voiceless stop in Spanish, the pattern is of the F1 peaking in the middle of the vowel, while falling at the beginning and end (Figure 12). This drop in F1 is steeper towards the end of the vowel. All three places of articulation converge at the end of the vowel as well. Vowels followed by /t/ and /k/ are fairly close to each other throughout the vowel, with vowels followed by /p/ having a lower F1. Place of articulation is significant at 20%, although vowels followed by /t/ do not present a difference from vowels preceded by /k/, and at 50%, although only vowels followed by /p/ and /k/. There is no significant effect of timepoint or interaction between factors.

4.3 VOT

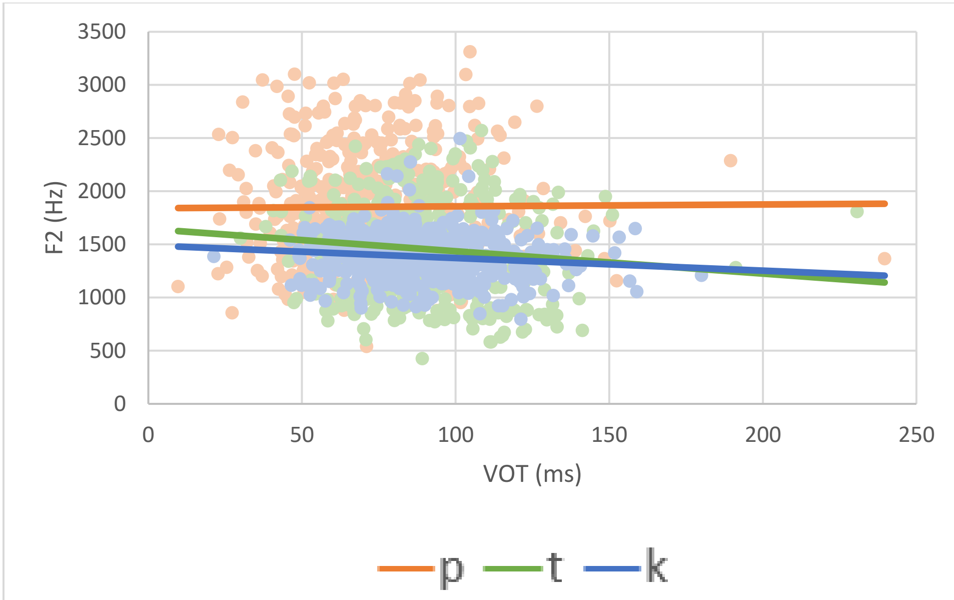

Figure 13. Average F2 of English vowels at 20% vowel duration by VOT and preceding voiceless stop.

For VOT, we can only look at general patterns because vowel quality will confound with differences in F2 caused by the preceding consonant. This data is from both timepoints-there was no significant different in VOT in either language between the beginning and the end of the semester. Figure 13 looks at F2 at 20% of vowel duration in English. In general, vowels preceded by /p/ have an F2 that rises slightly as VOT increases, while the F2 for vowels preceded by /t/ and /k/ falls as VOT lengthens. The overall pattern is that differences in F2 are greater as the VOT is longer for /p/, vs. /k/ and /t/.

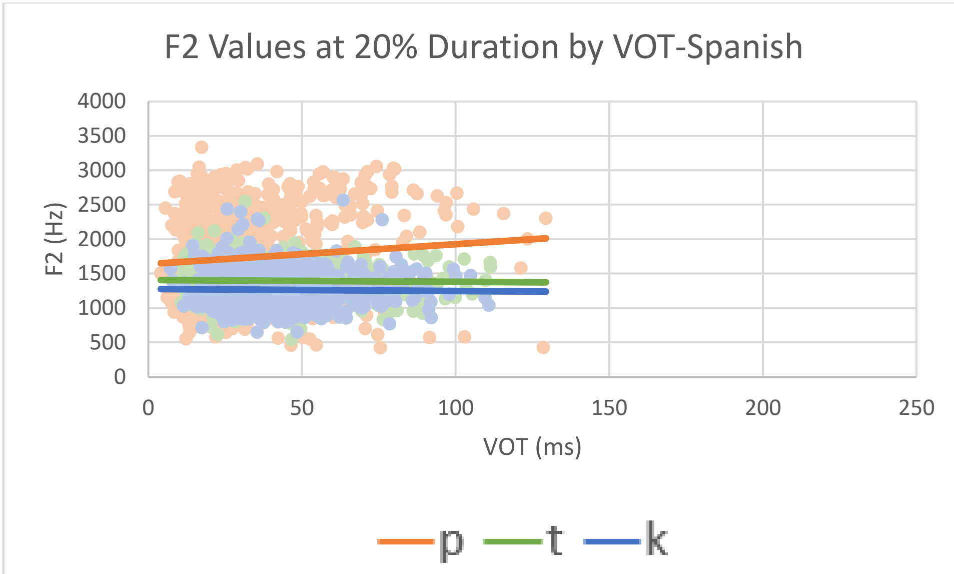

Figure 14. Average F2 of Spanish vowels at 20% duration by VOT and preceding voiceless stop.

In Spanish, there is a similar difference between vowels preceded by /t/ and /k/ and vowels preceded by /p/ (Figure 14). However, vowels preceded by /p/ increase more steeply as the VOT increases, while vowels preceded by /t/ and /k/ only decrease very slightly as VOT increases. The resulting trend is that differences in F2 in Spanish between /p/ and /t, k/ are greater as the VOT is longer.

5. Discussion

5.1 Differences between English and Spanish

	Preceding (English)	Preceding (Spanish)	Following (Spanish)
F2 of back vowels	p=k<t, then t<k<p	p<k<t	p=t=k
F2 of central vowels	N/A	p<t=k	p<k<t
F1 of mid vowels	p<k<t	p=t=k	p=t=k
F1 of low vowels	p<k	p<k, p=t, t=k	p<t=k

Table 3. Summary of Results: Vowel formants tendencies by adjacent voiceless stop.

Table 3 shows the main results. Each cell represents the formant values, from smallest to largest, based on the adjacent voiceless stop. These measurements represent the trend across the entire vowel, since the ordering tended not to change over the duration of the vowel. In English, there was insufficient data for central vowels. In English, there are some similarities across vowel groupings, but the results are not completely consistent. Mid and low vowels showed that vowels preceded by /p/ had a lower F1 than vowels preceded by /k/. The back vowels were less consistent, with a lot of change over the duration of the vowel, something that we do not see very often in any of the other vowel groupings.

For vowels preceded by voiceless stops in Spanish, we find overall that vowels preceded by /p/ have a lower F2 than vowels preceded by /t/ or /k/. The results are less clear for F1, with fewer differences depending on preceding stops. These findings for F1 do not depict an obvious coarticulatory pattern for Spanish vowels.

In looking at vowels followed by voiceless stops in Spanish, there are several different patterns. For the F2 of central vowels, vowels followed by /p/ had the lowest F2, then vowels followed by /k/, then vowels followed by /t/. But back vowels had equivalent F2s, regardless of their following stop. Therefore, there are only apparent coarticulatory effects of following stops on F2 for central vowels. For F1, vowels followed by /p/ had a smaller F1 for low vowels, but there were no significant differences between following consonants for mid vowels.

In summary, we see the strongest and most consistent effect of place of articulation on coarticulation in English, from preceding stops. The most general overall trend we can see in the data is that vowels preceded by /p/ tend to have lower formant values than vowels preceded by other consonants. When we do see robust differences in coarticulation from place of articulation, mostly in English, they patterned as p<k<t, regardless of F1 or F2, preceding or following.

Returning to the models of coarticulation discussed in Section 2, we do not find strong adherence to either one. The predictions from these models are repeated here in Table 4. The feature spreading model only makes predictions for anticipatory coarticulation, based on the concept that segments “look-ahead” to following segments. The proposed lack of carryover coarticulation was not observed in the data. The feature-spreading model bases its predictions for F1 on jaw height, and F2 on place of articulation. The predicted ordering of F1 and F2 values was also not observed.

The degree of articulatory constraint model predicts differences between carryover and anticipatory coarticulation, with more anticipatory coarticulation from more constrained sounds. The p<k<t prediction for F1 anticipatory coarticulation and F2 carryover coarticulation is borne out in the data, although the differences between carryover and anticipatory coarticulation are not.

Overall, the degree of articulatory constraint model is more applicable to these data. Although its predictions do not entirely line up with the observed data, none of the predictions of the feature-spreading model held. These results suggest some credence for the DAC model. The problems with this model seemed to arise in the t<k<p prediction that was never observed. Possibly, carryover and anticipatory coarticulation are not differentiated based on articulatory constraint. Given that differences between vowels adjacent to /t/ and /k/ were not as pronounced in comparison to vowels adjacent to /p/, other factors besides articulatory constraint could be impacting coarticulation to produce the observed results.

	F1		F2	
	Carryover (preceding stop)	Anticipatory (following stop)	Carryover (preceding stop)	Anticipatory (following stop)
Feature Spreading	p=t=k	t<k<p	p=t=k	k<t<p
Degree of Articulatory Constraint	t<k<p	p<k<t	p<k<t	t<k<p

Table 4. Summary of predictions from the coarticulatory models considered in this study.

The predictions shown in Table 4 are based on articulatory constraints, and are not language-specific. As discussed in Section 2, there can be language-specific effects of coarticulation, such as the language-specific nasal coarticulation discussed by Solé and Ohala. Although English results were more pronounced than in Spanish, the patterning was relatively similar. The major difference appeared to be in degree of coarticulation. This might be because there is not a phonological/phonetic distinction dividing English and Spanish, as there was in nasal coarticulation. This language-specific distinction might result in more differentiated coarticulation than we observe with these data.

5.2 Differences across the semester

Overall, we saw remarkably few differences in coarticulation from the beginning of the semester to the end. One potential explanation is that coarticulation did not change very much because there is not much change in stop or vowel production between the beginning and the end of the semester. One example of this lack of change is VOT. At the beginning of the semester, VOT was already significantly shorter in Spanish than in English, and there were no significant differences in VOT by the end of the semester. Although VOT is shorter in Spanish at the beginning of the semester, it is not at native-like levels. It could be that one semester of pronunciation training is not sufficient to shift pronunciation very much.

5.3 VOT

The impact of VOT was somewhat surprising. Although we do not see much of a difference between English and Spanish in the F1 and F2 values, we do see language-specific differences in coarticulation based on VOT. As VOT increased, the differences between vowels preceded by /p/ and vowels preceded by /t/ and /k/ increased, in both languages. The exact pattern, however, was slightly different between languages. In Spanish, the F2 of vowels preceded by /p/ stayed relatively constant, while the F2 of vowels preceded by /t/ and /k/ decreased. In English, the F2 of vowels preceded by /p/ increased, while the F2 of vowels preceded by /t/ and /k/ remained constant. But the overall effect was the same: as VOT increased, formant differences based on place of articulation increased. This was unexpected, because I had predicted that more distance from the closure would decrease coarticulation. This effect was predicted to occur because articulators would have more time to move between gestures. However, this does not seem to be in the case.

The effect cannot solely be a function of VOT, however, because we see different patterns in English and Spanish, even when the VOTs are at the same length. The pattern does not become more English-like once the VOT in Spanish approaches native English lengths. In general, though, we saw stronger and more consistent coarticulation in English than in Spanish, and perhaps the longer VOT is a factor in this effect, i.e. a more English-like VOT in Spanish correlates with more English patterns of coarticulation overall. Alternatively, perhaps there is more coarticulation with a longer VOT universally. However, it is unclear what the mechanism behind this finding is.

6. Conclusion and Future Directions

This study found some differences between L1 English and L2 Spanish in voiceless stop-vowel coarticulatory patterns, more precisely coarticulation was stronger in L1 English. It did not find much difference between the beginning and end of the semester of pronunciation training. It found an effect of VOT, with a longer VOT resulting in a greater coarticulatory effect.

The effect of VOT was a very interesting finding of this study. This is an area that would benefit from further study, especially looking at more specific VOT lengths, and the breakdown of VOT among different vowel groups. It would be beneficial to repeat this study with more comprehensive stimuli, in order to look at the effect of coarticulation overall, rather than breaking it down by vowel classes. The necessity of looking at each vowel group separately broke the large number of tokens down into very small groups, that made it very hard to find a clear picture of effects. It would also be important to analyze more data for some of the vowel-consonant combinations that were missing from the current study, especially to explore the effect of following consonants on vowels in English.

Because there was not very much change over the course of the semester, it would be interesting to look at changes in coarticulation over a wider span of learning, perhaps beginning with introductory Spanish, where students are perhaps unaware of the differences in pronunciation of voiceless stops. Finally, it would be interesting to compare these results with coarticulation in native Spanish speakers. This would help untangle some of the results that we see, such as the impact of VOT.

Appendix A-Statistical Tables

Best fit models from linear regression

BACK VOWELS

Table 1. Best-fit model for F2 at 20% for back vowels and preceding consonant dataset

	Estimate	Std. Error	df	t value	P-value
(Intercept)	1337.45	20.11	104.62	66.500	< 2e-16
Place of articulation (reference level: /k/)					
/p/	-21.46	28.54	1584.16	-0.752	0.452312
/t/	232.91	22.01	1583.75	10.584	< 2e-16
Language (reference level: English)					
Spanish	-181.26	16.91	1629.39	-10.721	< 2e-16
Place of articulation*Language					
/p/:Spanish	-125.19	34.58	1584.17	-3.621	0.000303
/t/:Spanish	-111.27	27.32	1583.82	-4.073	4.87e-05

Table 2. Best-fit model for F2 at 50% for back vowels and preceding consonant dataset

	Estimate	Std. Error	df	t value	P value
(Intercept)	1271.87	19.79	99.90	64.266	< 2e-16
Place of articulation (reference level: /k/)					
/p/	246.52	27.83	1582.63	8.858	< 2e-16
/t/	227.37	21.46	1582.21	10.597	< 2e-16
Language (reference level: English)					
Spanish	-171.29	16.49	1629.19	-10.387	< 2e-16
Place of articulation*Language					
/p/:Spanish	-352.42	33.71	1582.64	-10.453	< 2e-16
/t/:Spanish	-202.61	26.64	1582.29	-7.606	4.81e-14

Table 3. Best-fit model for F2 at 80% for back vowels and preceding consonant dataset

	Estimate	Std. Error	df	t value	P value
(Intercept)	1325.05	21.38	175.13	61.962	< 2e-16
Place of articulation (reference level: /k/)					
/p/	428.05	38.09	1582.95	11.239	< 2e-16
/t/	199.32	29.36	1582.25	6.788	1.61e-11
Language (reference level: English)					
Spanish	-202.86	22.29	1625.06	-9.102	< 2e-16
Place of articulation*Language					
/p/:Spanish	-465.84	46.14	1583.14	-10.097	< 2e-16
/t/:Spanish	-231.66	36.45	1582.51	-6.355	2.72e-10

Table 4. Best-fit model for F2 at 80% for back vowels and following consonant dataset

	Estimate	Std. Error	df	t value	P value
(Intercept)	1088.15	26.19	80.59	41.55	<2e-16
Timepoint (reference level: T1)					
T2	-80.58	32.11	244.55	-2.51	0.0127

CENTRAL VOWELS

Table 5. Best-fit model for F2 at 20% for central vowels and preceding consonant dataset

	Estimate	Std. Error	df	t value	P value
(Intercept)	1466.28	23.39	73.62	62.682	< 2e-16
Place of articulation (reference level: /k/)					
/p/	-154.82	14.51	839.07	-10.669	< 2e-16
/t/	16.72	14.58	839.12	1.147	0.251544
Language (reference level: English)					
Spanish	59.87	17.25	880.24	3.472	0.000542

Table 6. Best-fit model for F2 at 50% for central vowels and preceding consonant dataset

	Estimate	Std. Error	df	t value	P value
(Intercept)	1521.083	24.426	78.026	62.273	< 2e-16
Place of articulation (reference level: /k/)					
/p/	-35.992	15.844	839.995	-2.272	0.02336
/t/	9.301	15.915	840.052	0.584	0.55911
Language (reference level: English)					
Spanish	-61.144	18.792	881.798	-3.254	0.00118

Table 7. Best-fit model for F2 at 80% for central vowels and preceding consonant dataset

	Estimate	Std. Error	df	t value	P value
(Intercept)	1663.281	25.623	94.894	64.913	<2e-16
Place of articulation (reference level: /k/)					
/p/	15.686	19.596	839.787	0.800	0.424
/t/	-7.618	19.684	839.883	-0.387	0.699
Language (reference level: English)					
Spanish	-224.694	23.036	885.540	-9.754	<2e-16

Table 8. Best-fit model for F2 at 20% for central vowels and following consonant dataset

	Estimate	Std. Error	df	t value	P value
(Intercept)	1492.296	32.612	78.138	45.759	<2e-16
Place of articulation (reference level: /k/)					
/p/	9.697	22.738	403.366	0.426	0.6700
/t/	-94.208	29.643	403.596	-3.178	0.0016

Table 9. Best-fit model for F2 at 80% for central vowels and following consonant dataset

	Estimate	Std. Error	df	t value	P value
(Intercept)	1455.98	32.38	100.56	44.962	< 2e-16
Place of articulation (reference level: /k/)					
/p/	-99.25	25.33	403.26	-3.919	0.000104
/t/	62.95	33.02	403.60	1.907	0.057285

LOW VOWELS

Table 10. Best-fit model for F1 at 20% for low vowels and preceding consonant dataset

	Estimate	Std. Error	df	t value	P value
(Intercept)	851.57	15.49	87.40	54.973	< 2e-16
Language (reference level: English)					
Spanish	-68.88	10.34	882.93	-6.661	4.77e-11

Table 11. Best-fit model for F1 at 50% for low vowels and preceding consonant dataset

	Estimate	Std. Error	df	t value	P value
(Intercept)	883.15	16.40	90.00	53.846	< 2e-16
Language (reference level: English)					
Spanish	-77.89	11.35	882.55	-6.862	1.28e-11

Table 12. Best-fit model for F1 at 80% for low vowels and preceding consonant dataset

	Estimate	Std. Error	df	t value	P value
(Intercept)	810.51	24.38	283.53	33.244	< 2e-16
Place of articulation (reference level: /t/)					
/k/	25.36	16.40	841.32	1.546	0.12245
/p/	36.10	15.49	841.54	2.330	0.02003
Language (reference level: English)					
Spanish	-52.99	17.67	879.17	-2.999	0.00279

Table 13. Best-fit model for F1 at 20% for low vowels and following consonant dataset

	Estimate	Std. Error	df	t value	P value
(Intercept)	789.31	18.17	37.63	43.436	< 2e-16
Place of articulation (reference level: /p/)					
/k/	32.67	14.70	404.34	2.223	0.026771
/t/	51.92	15.13	404.98	3.431	0.000664

Table 14. Best-fit model for F1 at 50% for low vowels and following consonant dataset

	Estimate	Std. Error	df	t value	P value
(Intercept)	862.08	20.32	68.25	42.416	< 2e-16
Place of articulation (reference level: /k/)					
/p/	-40.72	13.10	404.25	-3.108	0.00201
/t/	-17.09	17.08	404.44	-1.000	0.31782

MID VOWELS

Table 15. Best-fit model for F1 at 20% for mid vowels and preceding consonant dataset

	Estimate	Std. Error	df	t value	P value
(Intercept)	687.611	10.244	74.885	67.125	< 2e-16
Place of articulation (reference level: /k/)					
/p/	-71.141	7.973	2034.335	-8.923	< 2e-16
/t/	89.668	9.115	2034.319	9.838	< 2e-16
Language (reference level: English)					
Spanish	-131.195	6.792	2068.915	-19.315	< 2e-16
Place of articulation*Language					
/p/:Spanish	62.643	10.377	2034.364	6.037	1.86e-09
/t/:Spanish	-90.405	11.471	2034.361	-7.881	5.23e-15

Table 16. Best-fit model for F1 at 50% for mid vowels and preceding consonant dataset

	Estimate	Std. Error	df	t value	P value
(Intercept)	655.519	10.422	93.724	62.900	< 2e-16
Place of articulation (reference level: /k/)					
/p/	-56.408	10.072	2034.200	-5.601	2.43e-08
/t/	136.676	11.514	2034.179	11.870	< 2e-16
Language (reference level: English)					
Spanish	-104.720	8.546	2078.887	-12.253	< 2e-16
Place of articulation*Language					
/p/:Spanish	41.101	13.109	2034.267	3.135	0.00174
/t/:Spanish	-143.371	14.490	2034.265	-9.894	< 2e-16

Table 17. Best-fit model for F1 at 80% for mid vowels and preceding consonant dataset

	Estimate	Std. Error	df	t value	P value
(Intercept)	612.732	12.031	156.699	50.930	< 2e-16
Place of articulation (reference level: /k/)					
/p/	-47.714	13.272	2035.346	-3.595	0.000332
/t/	152.387	15.173	2035.329	10.044	< 2e-16
Language (reference level: English)					
Spanish	-97.167	11.201	2081.792	-8.675	< 2e-16
Timepoint (reference level: T1)					
T2	-22.983	7.441	2081.389	-3.089	0.002036
Place of articulation*Language					
/p/:Spanish	34.001	17.274	2035.490	1.968	0.049170
/t/:Spanish	-162.613	19.095	2035.498	-8.516	< 2e-16

Table 18. Best-fit model for F1 at 50% for mid vowels and following consonant dataset

	Estimate	Std. Error	df	t value	P value
(Intercept)	544.069	10.887	46.288	49.976	<2e-16
Timepoint (reference level: T1)					
T2	-17.545	7.755	366.107	-2.262	0.0243

Table 19. Best-fit model for F1 at 80% for mid vowels and following consonant dataset

	Estimate	Std. Error	df	t value	P value
(Intercept)	495.55	11.01	66.63	45.017	<2e-16
Timepoint (reference level: T1)					
T2	-29.89	11.66	373.59	-2.563	0.0108

Works Cited

Boersma, P., & Weenink, D. (2005). Praat. *Doing phonetics by computer (Version 5.1)*.

Farnetani, E., & Recasens, D. (1997). Coarticulation and connected speech processes. *The handbook of phonetic sciences*, 371-404.

Fowler, C. A., & Saltzman, E. (1993). Coordination and coarticulation in speech production. *Language and speech*, *36*(2-3), 171-195.

Hammarberg, R. (1976). The metaphysics of coarticulation. *Journal of Phonetics*, *4*(4), 353-363.

Hualde, J. I. (2005). *The sounds of Spanish with audio CD*. Cambridge University Press.

Keating, P. A. (1988). Underspecification in phonetics. *Phonology*, *5*(2), 275-292.

Menke, M. R., & Face, T. L. (2010). Second Language Spanish Vowel Production: An Acoustic Analysis. *Studies in Hispanic & Lusophone Linguistics*, *3*(1).

Recasens, D. (1987). An acoustic analysis of V-to-C and V-to-V coarticulatory effects in Catalan and Spanish VCV sequences. *Journal of Phonetics*, *15*(4), 299-312.

Recasens, D. (2012). A cross-language acoustic study of initial and final allophones of/l. *Speech Communication*, *54*(3), 368-383.

Recasens, D. (2012). A study of jaw coarticulatory resistance and aggressiveness for Catalan consonants and vowels. *The Journal of the Acoustical Society of America*, *132*(1), 412-420.

Recasens, D., & Farnetani, E. (1990). Articulatory and acoustic properties of different allophones of/l/in American English, Catalan and Italian. In *First International Conference on Spoken Language Processing*.

Recasens, D., Pallarès, M. D., & Fontdevila, J. (1997). A model of lingual coarticulation based on articulatory constraints. *The Journal of the Acoustical Society of America*, *102*(1), 544-561.

Sole, M. J., & Ohala, J. J. (1991). Differentiating between phonetic and phonological processes: The case of nasalization. *Actes du XIIème Congrès International des Sciences Phonétiques, Aix en Provence, 3*, 110-113.

Solon, M. (2018). Acquisition of articulatory control or language-specific coarticulatory patterns?: Evidence from the production of laterals in second-language Spanish. In *Contemporary Trends in Hispanic and Lusophone Linguistics* (pp. 213-236). John Benjamins.

Stevens, K. N., & House, A. S. (1963). Perturbation of vowel articulations by consonantal context: An acoustical study. *Journal of speech and hearing research*, *6*(2), 111-128.

Volenec, V. (2015). Coarticulation. In *Phonetics: Fundamentals, Potential Applications and Role in Communicative Disorders*. Nova Science Publishers.

Wanjema, S., Carmichael, K., Walker, A., & Campbell-Kibler, K. (2013). The Ohiospeaks project: Engaging undergraduates in sociolinguistic research. *American speech*, *88*(2), 223-235.